### **RAW SEQUENCE LISTING**

- The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 1573,600A
Source: 1700
Date Processed by STIC: 8/23/06

ENTERED



**IFWO** 

RAW SEQUENCE LISTING DATE: 08/23/2006
PATENT APPLICATION: US/10/573.600A TIME: 10:23:13

Input Set: N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

```
1 <110> APPLICANT: The Trustees of the University of Pennsylvania
        Wilson, James M.
 2
        Gao, Guangping
 3
 4
        Alvira, Mauricio R.
        Vandenberghe, Luk H.
 6 <120> TITLE OF INVENTION: Adeno-Associated Virus (AAV) Clades, Sequences, Vectors
 7
         Containing Same, and Uses Therefor
 8 <130> FILE REFERENCE: UPN-P3230PCT
 9 <140> CURRENT APPLICATION NUMBER: US/10/573,600A
10 <141> CURRENT FILING DATE: 2006-03-24
11 <150> PRIOR APPLICATION NUMBER: US 60/508,226
12 <151> PRIOR FILING DATE: 2003-09-30
13 <150> PRIOR APPLICATION NUMBER: US 60/566,546
14 <151> PRIOR FILING DATE: 2004-04-29
15 <160> NUMBER OF SEQ ID NOS: 236
16 <170> SOFTWARE: PatentIn version 3.3
18 <210> SEO ID NO: 1
19 <211> LENGTH: 2211
20 <212> TYPE: DNA
21 <213> ORGANISM: Unknown
22 <220> FEATURE:
23 <223> OTHER INFORMATION: adeno-associated virus, clone hu.31
24 <400> SEQUENCE: 1
25
                                                                                60
         atggctgccg atggttatct tccagattgg ctcgaggaca accttagtga aggaattcgc
         gagtggtggg ctttgaaacc tggagcccct caacccaagg caaatcaaca acatcaagac
26
                                                                               120
27
         aacgctcgag gtcttgtgct tccgggttac aaataccttg gacccggcaa cggactcgac
                                                                               180
28
                                                                               240
         aagggggage eggteaaege ageagaegeg geggeeeteg ageaegaeaa ggeetaegae
29
                                                                               300
         cagcagetea aggeeggaga caaccegtae etcaagtaca accaegeega egeegagtte
30
         caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag
                                                                               360
31
        gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct
                                                                               420
32
        gqaaaqaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc
                                                                               480
33
         aaatcqqqtq cacaqcccqc taaaaaqaqa ctcaatttcg gtcaqactgg cgacacagag
                                                                               540
                                                                               600
34
         teagteecag acceteaace aateggagaa ceteeegeag ecceeteagg tgtgggatet
35
                                                                               660
        cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga
36
        gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc
                                                                               720
37
        accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc
                                                                               780
38
        tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcacccc
                                                                               840
        tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga
39
                                                                               900
40
        ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt
                                                                               960
        caggicaaag aggitacgga caacaatgga gicaagacca tcgccaataa ccitaccagc
41
                                                                              1020
42
                                                                              1080
        acggtccagg tcttcacgga ctcagactat cagctcccgt acgtgctcgg gtcggctcac
43
        gagggetgee teeegeegtt eecageggae gtttteatga tteeteagta egggtatetg
                                                                              1140
        acgettaatg atggaageca ggeegtgggt egttegteet tttactgeet ggaatattte
                                                                              1200
44
```

### RAW SEQUENCE LISTING DATE: 08/23/2006 PATENT APPLICATION: US/10/573,600A TIME: 10:23:13

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

```
45
                                                                               1260
         ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta
                                                                               1320
46
         cetttecata geagetaege teacageeaa ageetggaee gaetaatgaa tecaeteate
47
         gaccaatact tgtactatct ctcaaagact attaacggtt ctggacagaa tcaacaaacg
                                                                               1380
48
         ctaaaattca gtgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct
                                                                               1440
49
         ggacccaget accgacaaca acgtgtetea accaetgtga etcaaaacaa caacagegaa
                                                                               1500
50
         tttqcttqqc ctqqaqcttc ttcttqqqct ctcaatqqac gtaataqctt gatgaatcct
                                                                               1560
51
         ggacctgcta tggccagcca caaagaagga gaggaccgtt tctttccttt gtctggatct
                                                                               1620
52
         ttaatttttg gcaaacaagg aactggaaga gacaacgtgg atgcggacaa agtcatgata
                                                                               1680
53
                                                                               1740
         accaacgaag aagaaattaa aactactaac ccggtagcaa cggagtccta tggacaagtg
         gccacaaacc accagagtgc ccaagcacag gcgcagaccg gctgggttca aaaccaagga
                                                                               1800
54
55
         atacttccgg gtatggtttg gcaggacaga gatgtgtacc tgcaaggacc catttgggcc
                                                                               1860
56
         aaaattcctc acacggacgg caactttcac ccttctccgc tgatgggagg gtttggaatg
                                                                               1920
57
         aagcaccege cteeteagat eeteateaaa aacacacetg tacetgegga teeteeaacg
                                                                               1980
58
         gccttcaaca aggacaagct gaactctttc atcacccagt attctactgg ccaagtcagc
                                                                               2040
59
         gtggagatcg agtgggagct gcagaaggaa aacagcaagc gctggaaccc ggagatccag
                                                                               2100
         tacacttcca actattacaa gtctaataat gttgaatttg ctgttaatac tgaaggtgta
                                                                               2160
60
         tatagtgaac cccgccccat tggcaccaga tacctgactc gtaatctgta a
                                                                               2211
63 <210> SEQ ID NO: 2
64 <211> LENGTH: 2211
65 <212> TYPE: DNA
66 <213> ORGANISM: Unknown
67 <220> FEATURE:
68 <223> OTHER INFORMATION: new AAV serotype, clone hu.32
69 <400> SEQUENCE: 2
70
         atggctgccg atggttatct tccagattgg ctcgaggaca ctctctctga aggaataaga
                                                                                 60
71
         cagtggtgga agctcaaacc tggcccacca ccaccaaagc ccgcagagcg gcataaggac
                                                                                120
72
                                                                                180
         gacagcaggg gtcttgtgct tcctgggtac aagtacctcg gacccggcaa cggactcgac
73
                                                                                240
         aagggggage eggteaaege ageagaegeg geggeeeteg ageaegaeaa ggeetaegae
74
                                                                                300
         cagcagetea aggeeggaga caaccegtae etcaagtaca accaegeega egeegagtte
75
         caggagcggc tcaaagaaga tacgtctttt gggggcaacc tcgggcgagc agtcttccag
                                                                                360
76
                                                                                420
         gccaaaaaga ggcttcttga acctcttggt ctggttgagg aagcggctaa gacggctcct
                                                                                480
77
         ggaaagaaga ggcctgtaga gcagtctcct caggaaccgg actcctccgc gggtattggc
78
         aaatcgggtt cacagcccgc taaaaagaaa ctcaatttcg gtcagactgg cgacacagag
                                                                                540
79
         teagteeceg acceteaace aateggagaa ceteecgeag eeceeteagg tgtgggatet
                                                                                600
80
         cttacaatgg cttcaggtgg tggcgcacca gtggcagaca ataacgaagg tgccgatgga
                                                                                660
81
         gtgggtagtt cctcgggaaa ttggcattgc gattcccaat ggctggggga cagagtcatc
                                                                                720
82
         accaccagca cccgaacctg ggccctgccc acctacaaca atcacctcta caagcaaatc
                                                                                780
83
         tccaacagca catctggagg atcttcaaat gacaacgcct acttcggcta cagcacccc
                                                                                840
                                                                                900
84
         tgggggtatt ttgacttcaa cagattccac tgccacttct caccacgtga ctggcagcga
85
         ctcatcaaca acaactgggg attccggcct aagcgactca acttcaagct cttcaacatt
                                                                                960
86
         caggtcaaag aggttacgga caacaatgga gtcaagacca tcgccaataa ccttaccagc
                                                                               1020
87
         acggtccagg tettcacgga etcagaetat cagetceegt acgtgetegg gteggeteae
                                                                               1080
                                                                               1140
88
         gagggetgee teeegeegtt eeeageggae gtttteatga tteeteagta egggtatetg
89
         acgettaatg atgggageca ggeegtgggt egttegteet tttactgeet ggaatattte
                                                                               1200
90
         ccgtcgcaaa tgctaagaac gggtaacaac ttccagttca gctacgagtt tgagaacgta
                                                                               1260
                                                                               1320
         cetttecata geagetacge teacageeaa ageetggace gaetaatgaa tecaeteate
91
92
         qaccaatact tqtactatct ctcaaaqact attaacqqtt ctqqacaqaa tcaacaaacq
                                                                               1380
93
         ctaaaattca gcgtggccgg acccagcaac atggctgtcc agggaagaaa ctacatacct
                                                                               1440
94
         ggacccaget accgacaaca acgtgtetea accaetgtga etcaaaacaa caacagegaa
                                                                               1500
```

# RAW SEQUENCE LISTING DATE: 08/23/2006 PATENT APPLICATION: US/10/573,600A TIME: 10:23:13

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

95	tttgcttggc ctggagcttc	ttcttgggct	ctcaatggac	gtaatagctt g	gatgaatcct	1560
96	ggacctgcta tggccagcca					1620
97	ttaatttttg gcaaacaagg	aactggaaga	gacaacgtgg	atgcggacaa	agtcatgata	1680
98	accaacgaag aagaaattaa	aactactaac	ccggtagcaa	cggagtccta	tggacaagtg	1740
99	gccacaaacc accagagtgc	ccaagcacag	gcgcagaccg	gctgggttca	aaaccaagga	1800
100	atacttccgg gtatggtttg					1860
101	aaaattcctc acacggacgg					1920
102	aagcacccgc ctcctcagat					1980
103	gctttcaata aggacaagct					2040
104	gtggagattg agtgggagct					2100
105	tacacttcca actattacaa					2160
106	tatagtgaac cccgccccat	tggcaccaga	tacctgactc	gtaatctgta	a	2211
108 <	210> SEQ ID NO: 3					
109 <	211> LENGTH: 2211					
110 <	212> TYPE: DNA					
111 <	213> ORGANISM: Unknown					
112 <	220> FEATURE:					
113 <	223> OTHER INFORMATION: add	eno-associa	ted virus,	human clone	9	
114 <	<pre>&lt;400&gt; SEQUENCE: 3</pre>					
115	atggctgccg atggttatct	tccagattgg	ctcgaggaca	accttagtga	aggaattcgc	60
116	gagtggtggg ctttgaaacc	tggagcccct	caacccaagg	caaatcaaca	acatcaagac	120
117	aacgctcgag gtcttgtgct	tccgggttac	aaataccttg	gacccggcaa	cggactcgac	180
118	aagggggagc cggtcaacgc	agcagacgcg	gcggccctcg	agcacgacaa	ggcctacgac	240
119	cagcagctca aggccggaga	caacccgtac	ctcaagtaca	accacgccga	cgccgagttc	300
120	caggagcggc tcaaagaaga					360
121	gccaaaaaga ggcttcttga	acctcttggt	ctggttgagg	aagcggctaa	gacggctcct	420
122	ggaaagaaga ggcctgtaga	gcagtctcct	caggaaccgg	actcctccgc	gggtattggc	480
123	aaatcgggtg cacagcccgc	taaaaagaga	ctcaatttcg	gtcagactgg	cgacacagag	540
124	tcagtcccag accctcaacc	aatcggagaa	cctcccgcag	cccctcagg	tgtgggatct	600
125	cttacaatgg cttcaggtgg					660
126	gtgggtagtt cctcgggaaa	ttggcattgc	gattcccaat	ggctggggga	cagagtcatc	720
127	accaccagca cccgaacctg	ggccctgccc	acctacaaca	atcacctcta	caagcaaatc	780
128	tccaacagca catctggagg	atcttcaaat	gacaacgcct	acttcggcta	cagcaccccc	840
129	tgggggtatt ttgacttcaa	cagattccac	tgccacttct	caccacgtga	ctggcagcga	900
130	ctcatcaaca acaactgggg					960
131	caggtcaaag aggttacgga	caacaatgga	gtcaagacca	tcgccaataa	ccttaccagc	1020
132	acggtccagg tcttcacgga					1080
133	gagggctgcc tcccgccgtt					1140
134	acgcttaatg atggaagcca			-		1200
135	ccgtcgcaaa tgctaagaac					1260
136	cctttccata gcagctacgc					1320
137	gaccaatact tgtactatct	_			_	1380
138	ctaaaattca gtgtggccgg					1440
139	ggacccagct accgacaaca					1500
140	tttgcttggc ctggagcttc					1560
141	ggacctgcta tggccagcca					1620
142	ttaatttttg gcaaacaagg					1680
143	accaacgaag aagaaattaa					1740
144	gccacaaacc accagagtgc	ccaagcacag	gcgcagaccg	gctgggttca	aaaccaagga	1800

# RAW SEQUENCE LISTING DATE: 08/23/2006 PATENT APPLICATION: US/10/573,600A TIME: 10:23:13

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

	atacttccgg gtar aaaattcctc acac aagcacccgc ctcc gccttcaaca agga gtggagatcg agts tacacttcca acts tatagtgaac cccs SEQ ID NO: 4 LENGTH: 2217	cggacgg ctcagat acaagct gggagct attacaa	caactttcac cctcatcaaa gaactctttc gcagaaggaa gtctaataat	ccttctccgc aacacacctg atcacccagt aacagcaagc gttgaatttg	tgatgggagg tacctgcgga attctactgg gctggaaccc ctgttaatac	gtttggaatg tcctccaacg ccaagtcagc ggagatccag tgaaggtgta	1860 1920 1980 2040 2100 2160 2211		
	TYPE: DNA								
	ORGANISM: Unknown								
157 <220>	FEATURE:								
158 <223>	OTHER INFORMAT	ION: new	AAV seroty	pe, clone h	nu.17				
	SEQUENCE: 4								
160	atggctgccg atgg	gttatct	tccagattgg	ctcgaggaca	acctctctga	gggcattcgc	60		
161	gagtggtggg act						120		
162	gacggccggg gtc						180		
163	aagggggagc ccg						240		
164	cagcagctca aag						300		
165	caggagcgtc tgc						360		
166	gccaagaagc ggg						420		
167	ggaaagaaga gac				-		480		
168	ggcaagacag gcca						540		
169 170	gagtcagtgc ccg						600		
170	tctggtacaa tgg						660 720		
172	ggagtgggta gtt						780		
173	atcaccacca gca						840		
174	atctccaacg ggad ccctgggggt att					-	900		
175	cgactcatca acaa						960		
176	atccaggtca agga						1020		
177	agcacgattc aggi						1080		
178	caccagggct gcc						1140		
179	ctgactctga acaa						1200		
180	tttccttctc aaat						1260		
181	gtgccttttc acag						1320		
182	atcgaccagt acct						1380		
183	cagcagttgc tatt			_			1440		
184	ctacccgggc ccts						1500		
185	agcaactttg cttg						1560		
186	aatcccggtg tcgc	ctatggc	aacgcacaag	gacgacgaag	agcgattttt	tccatccagc	1620		
187	ggagtcttga tgtt	ttgggaa	acagggagct	ggaaaagaca	acgtggacta	tagcagcgtt	1680		
188	atgctaacca gtga	aggaaga	aatcaaaacc	accaacccag	tggccacaga	acagtacggc	1740		
189	gtggtggccg ataa	acctgca	acagcaaaac	gccgctccta	ttgtaggggc	cgtcaacagt	1800		
190	caaggagcct tacc						1860		
191	tgggccaaga ttcc						1920		
192	ggactgaaac acco						1980		
193	ccaactacct tcag						2040		
194	gtcagcgtgg aaat	tgaatg	ggagctgcag	aaagagaaca	gcaagcgctg	gaacccagag	2100		

#### RAW SEQUENCE LISTING

DATE: 08/23/2006 TIME: 10:23:13

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

PATENT APPLICATION: US/10/573,600A

195		attcagtata	cttccaacta	taacaaatct	gttaatgtgg	actttactgt	ggacactaat	2160			
196		ggtgtgtatt	cagagcctcg	ccccattggc	accagatacc	tgactcgtaa	tctgtaa	2217			
198	<210>	SEQ ID NO:					-				
199	<211>	LENGTH: 221	.7								
200	<212>	TYPE: DNA									
201	<213>	ORGANISM: Unknown									
202	<220>	FEATURE:									
203	<223>	OTHER INFORMATION: new AAV serotype, clone hu.6									
		SEQUENCE: 5									
205		atggctgccg	atggttatct	tccagattgg	ctcgaggaca	acctctctga	gggcattcgc	60			
206					aaacccaaag			120			
207					aagtacctcg			180			
208					gcggccctcg			240			
209					ctgcggtata			300			
210					gggggcaacc			360			
211					ctggttgagg			420			
212					cagcgttctc			480			
213					agactcaact			540			
214					gaaccccccg			600			
215					ccaatggcag			660			
216		ggagtgggta						720			
217					cccacctaca			780			
218					aacgacaaca		_	840			
219					cactgccact			900			
220					cccaagagac			960			
221					ggcaccaaga			1020			
222					taccagetee	_		1080			
223		caccagggct						1140			
224		ctgactctga						1200			
225		tttccttctc						1260			
226		gtgccttttc						1320			
227		atcgaccagt						1380			
228		cagcagttgc						1440			
229		ctacccgggc						1500			
230		agcaactttg						1560			
231		aatcccggtg						1620			
232		ggagtcttga						1680			
233		atgctaacca						1740			
234		gtggtggccg						1800			
235		caaggagcct						1860			
236		tgggccaaga	ttcctcacac	ggacggcaac	tttcatcctt	cgccgctgat	gggaggcttt	1920			
237		ggactgaaac	acccgcctcc	tcagatcctg	attaagaata	cacctgttcc	cgcggatcct	1980			
238		ccaactacct						2040			
239		gtcagcgtgg	aaattgaatg	ggagctgcag	aaagagaaca	gcaagcgctg	gaacccagag	2100			
240		attcagtata	cttccaacta	ctacaaatct	acaaatgtgg	actttgctgt	caatactgag	2160			
241		ggtacttatt		ccccattggc	acccgttacc	tcacccgtaa	cctgtaa	2217			
243	<210>	SEQ ID NO:	6			-					
244	<211>	LENGTH: 221	7								
245	<212>	TYPE: DNA									

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 08/23/2006
PATENT APPLICATION: US/10/573,600A TIME: 10:23:14

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

#### Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 6

VERIFICATION SUMMARY

DATE: 08/23/2006 TIME: 10:23:14

Input Set : N:\Crf4\08182006\J573600.raw
Output Set: N:\CRF4\08232006\J573600A.raw

PATENT APPLICATION: US/10/573,600A